# 4D Hybrid EnVar and Other DA Developments for the NCEP GFS:

## Potential for Next Generation Reanalysis

**Daryl Kleist[1]**
Rahul Mahajan[2], Catherine Thomas[1,2], Deng-Shun Chen[2,4]
Jeff Whitaker[3], Lili Lei[3], Phil Pegion[3]
John Derber[2], Andrew Collard[2], Yanqiu Zhu[2],Emily Liu[2], Russ Treadon[2]

[1]Univ. of Maryland-College Park, Dept. of Atmos. & Oceanic Science
[2]NOAA/NCEP/Environmental Modeling Center
[3]NOAA/OAR/Earth System Research Laboratory
[4]Central Weather Bureau and National Central University

**Reanalysis Technical Workshop**
**NCEP, 4-5 May 2015**

# Ensemble-Var methods: nomenclature

- ***En-4DVar***: Propagate ensemble $\mathbf{P}^b$ from one assimilation window to the next (updated using EnKF for example), replace static $\mathbf{P}^b$ with ensemble estimate of $\mathbf{P}^b$ at start of 4DVar window, $\mathbf{P}^b$ propagated with tangent linear model within window.

- ***4D-EnVar***: $\mathbf{P}^b$ at every time in the assimilation window comes from ensemble estimate (TLM no longer used).

- As above, with ***hybrid*** in name: $\mathbf{P}^b$ is a linear combination of static and ensemble components.

- ***3D-EnVar***: same as 4D ensemble Var, but $\mathbf{P}^b$ is assumed to be constant through the assimilation window (current NCEP implementation).

***Courtesy: Jeff Whitaker***

# Hybrid 4D-Ensemble-Var [H-4DEnVar]

The Hybrid EnVar cost function can be easily extended to 4D and include a static contribution

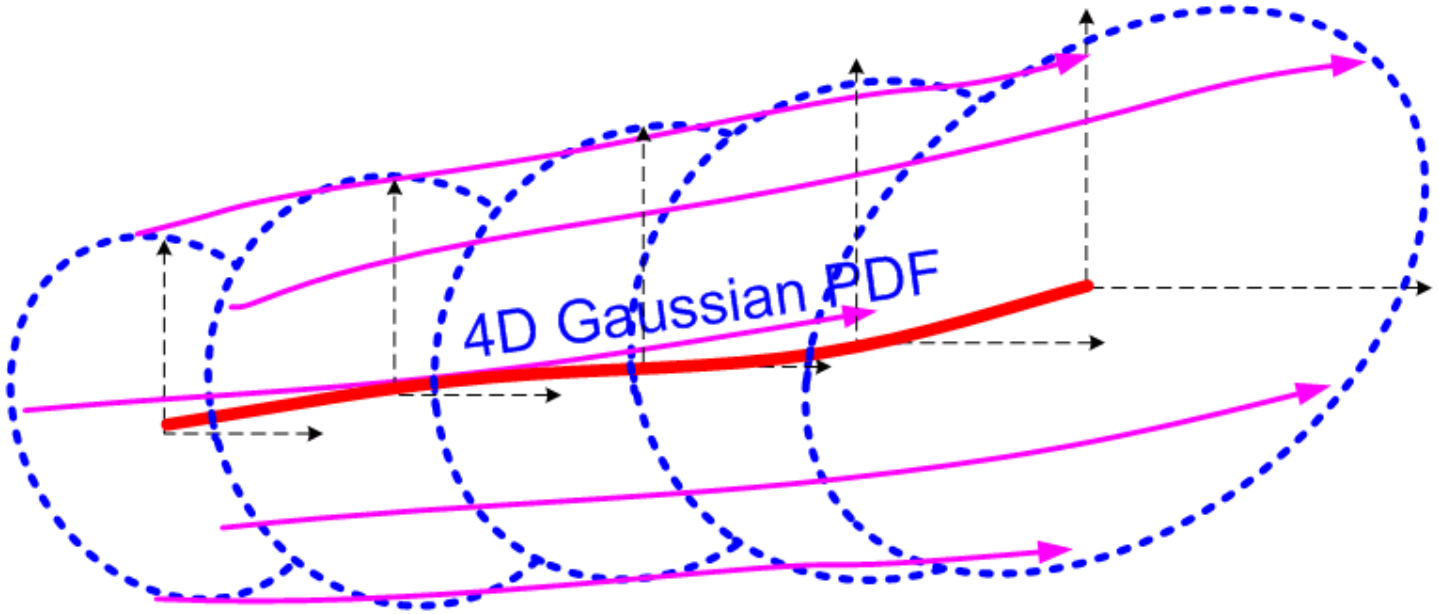
$$J\left(\mathbf{x}'_{\mathrm{f}},\boldsymbol{\alpha}\right)=\beta_{\mathrm{f}}\,\frac{1}{2}\left(\mathbf{x}'_{\mathrm{f}}\right)^{\mathrm{T}}\mathbf{B}_{\mathrm{f}}^{-1}\left(\mathbf{x}'_{\mathrm{f}}\right)+\beta_{\mathrm{e}}\frac{1}{2}\sum_{n=1}^{N}\left(\boldsymbol{\alpha}^{n}\right)^{\mathrm{T}}\mathbf{L}^{-1}\left(\boldsymbol{\alpha}^{n}\right)+$$
$$\frac{1}{2}\sum_{k=1}^{K}\left(\mathbf{H}_{k}\mathbf{x}'_{k}-\mathbf{y}'_{k}\right)^{\mathrm{T}}\mathbf{R}_{k}^{-1}\left(\mathbf{H}_{k}\mathbf{x}'_{k}-\mathbf{y}'_{k}\right)$$

Where the 4D increment is prescribed through linear combinations of the 4D ensemble perturbations plus static contribution

$$\mathbf{x}'_{k}=\mathbf{x}'_{\mathrm{f}}+\sum_{n=1}^{N}\left(\boldsymbol{\alpha}^{n}\circ\left(\mathbf{x}_{\mathrm{e}}\right)_{k}^{n}\right)$$

Here, the static contribution is considered time-invariant (i.e. from 3DVAR-FGAT). Weighting parameters exist just as in the other hybrid variants. No need for tangent linear/adjoint models.

# 4D EnVar




Trajectories of perturbations from ensemble mean

Full model evolves mean of PDF

Localised trajectories define 4D PDF of possible increments

4D analysis is a (localised) linear combination of nonlinear trajectories. It is not itself a trajectory.

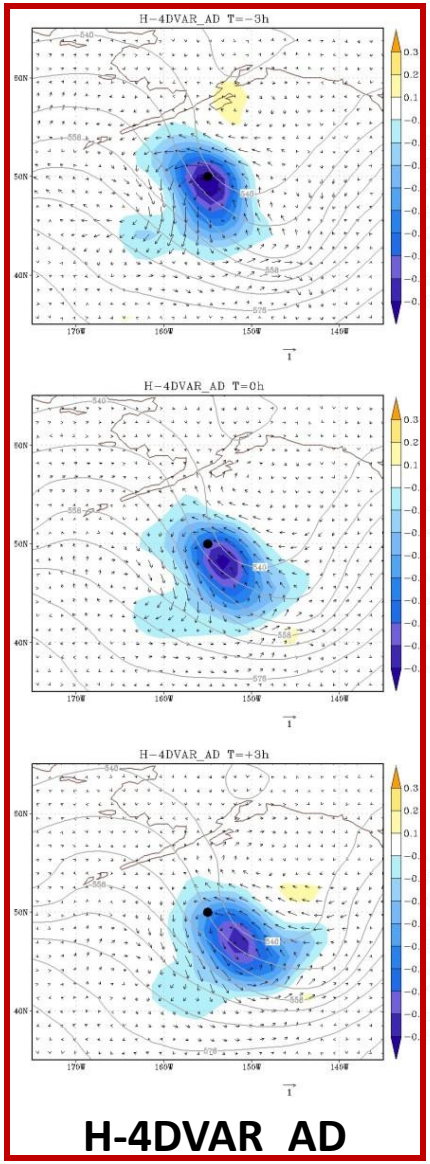
Courtesy: Andrew Lorenc

# Single Observation (-3h) Example From Kleist and Ide (2015)

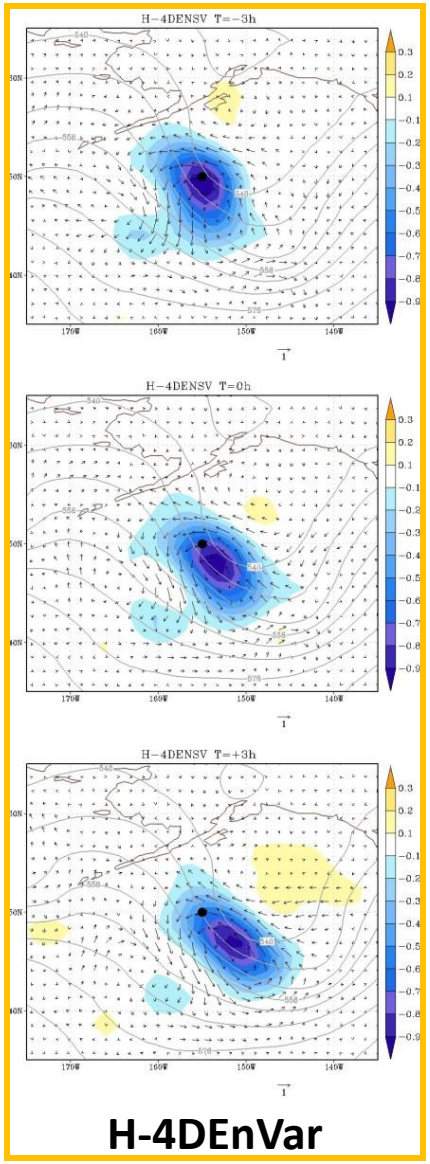
t=-3h

t=0h

t=+3h




H-4DVAR AD




H-4DEnVar

Solution at beginning of window same to within round-off (because observation is taken at that time, and same weighting parameters used)

Evolution of increment qualitatively similar between dynamic and ensemble specification

** Current linear and adjoint models in GSI are computationally unfeasible for use in 4DVAR other than simple single observation testing at low resolution

# GFS/GDAS Cycling
# Experiments with real observations

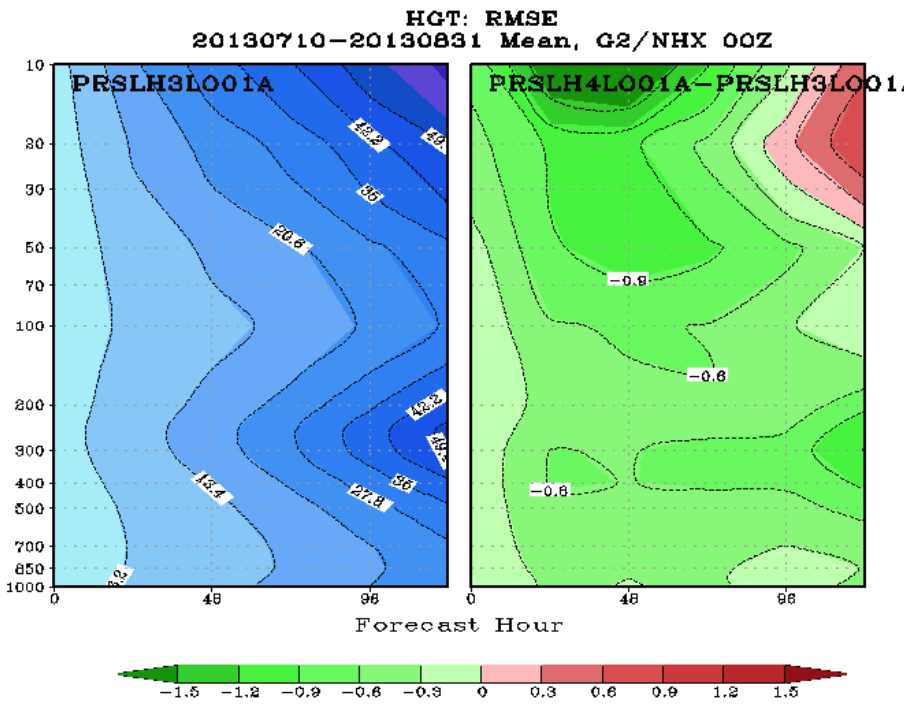
- Basic configuration
  - T670L64 Semi-Lagrangian GFS, operational observations, GFS/GDAS cycles

- **Hybrid 3D EnVar**
  - 80 member T254L46 ensemble with fully coupled (two-way) EnKF update
  - Incremental normal mode initialization (TLNMC) on total increment
  - 87.5% ensemble & 12.5% static
  - Multiplicative inflation and stochastic physics for EnKF perturbations

- **Hybrid 4D EnVar**
  - As in 3D Experiment, but extended to 4D
  - TLNMC on all time levels
  - Hourly TC relocation, O-G, binning of observations (not 3-hourly)

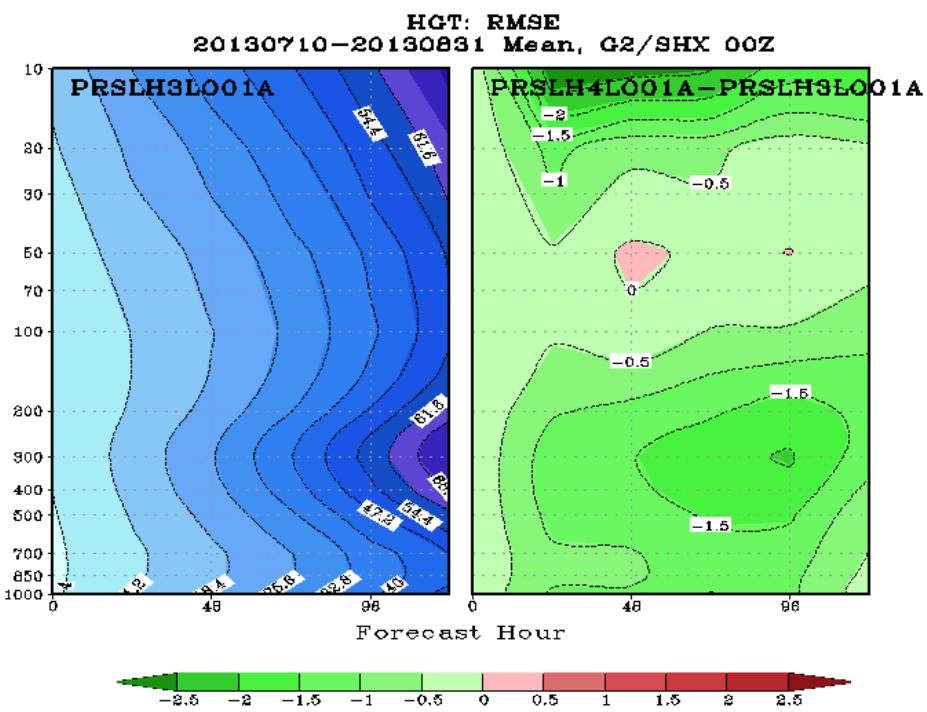
# 3D v 4D hybrid in SL GFS 20130710-20130831

**Hyb 4DEnVar-3DEnVar**




Northern Hemisphere Height RMSE and Difference

**Hyb 4DEnVar-3DEnVar**



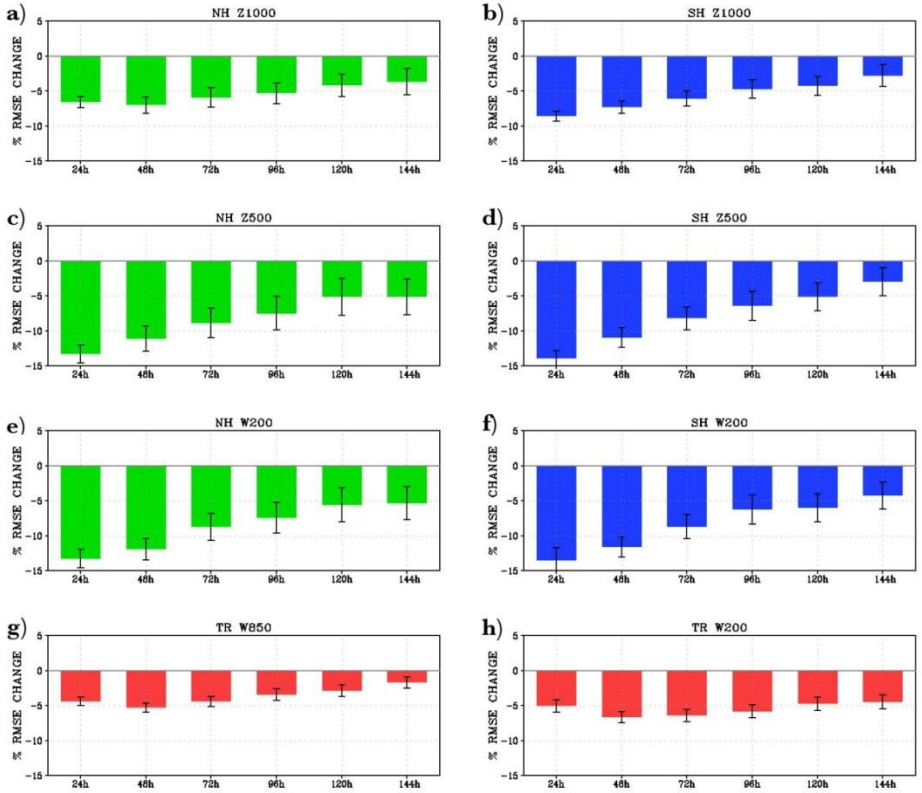

Southern Hemisphere Height RMSE and Difference

# Hybrid Assimilation Trials

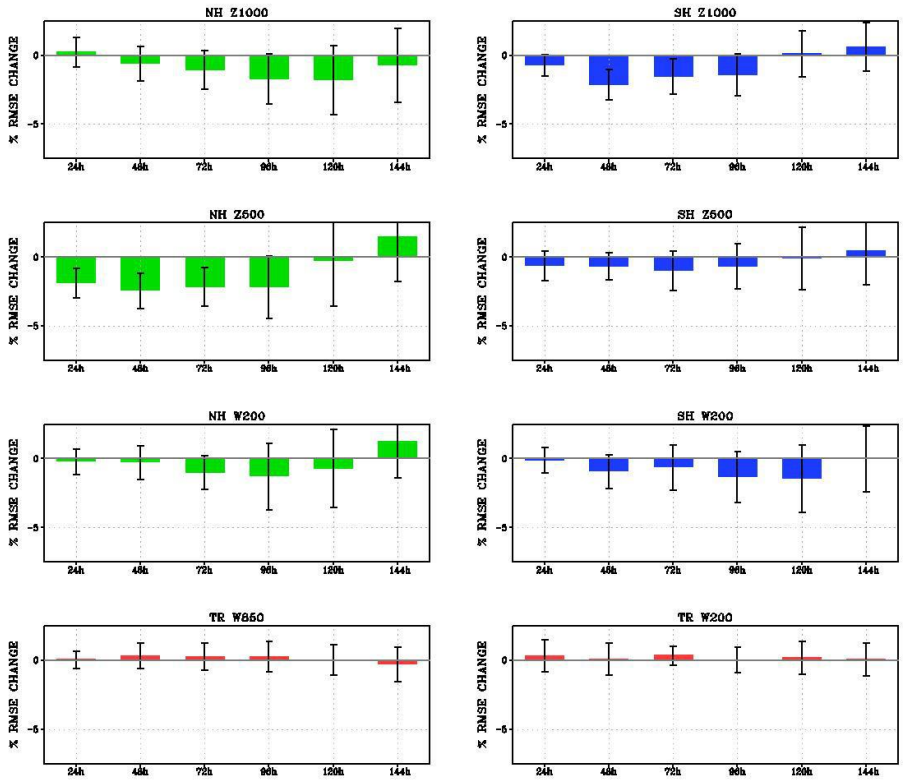
representative of many metrics/lead including tropical cyclone track

* Not an apples to apples comparison

**H 3DEnVar – 3DVAR**



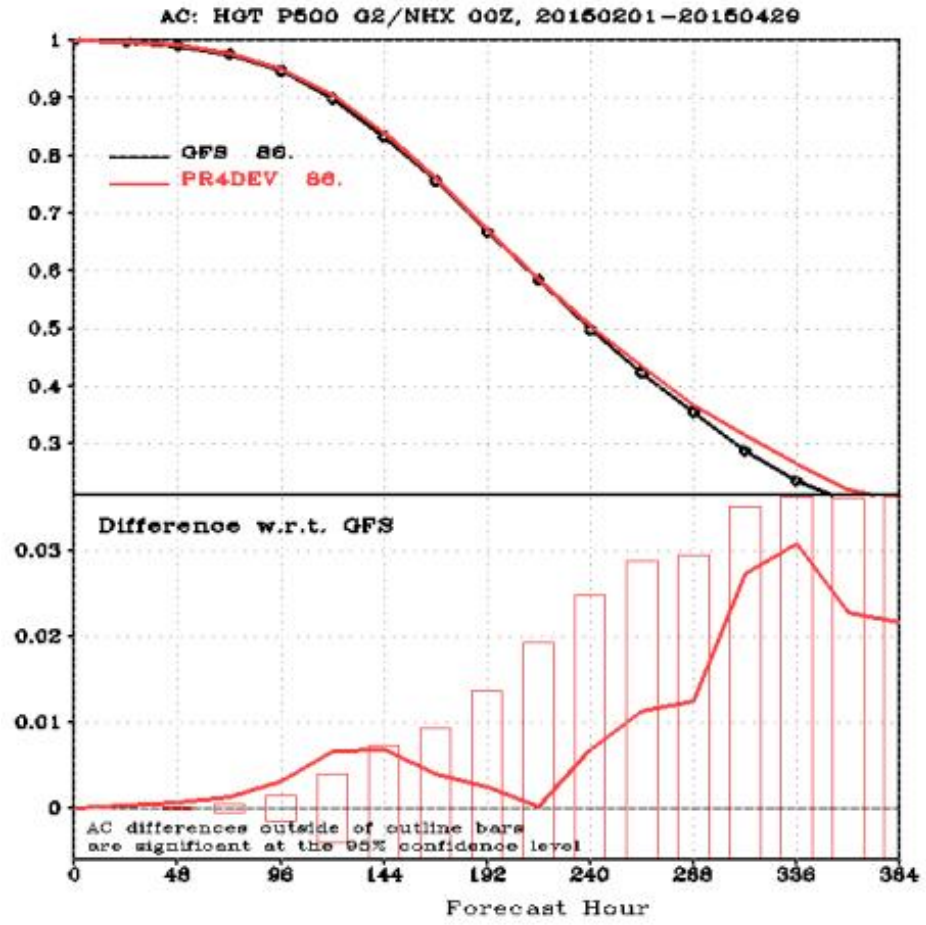

**H 4DEnVar – H 3DEnVar**

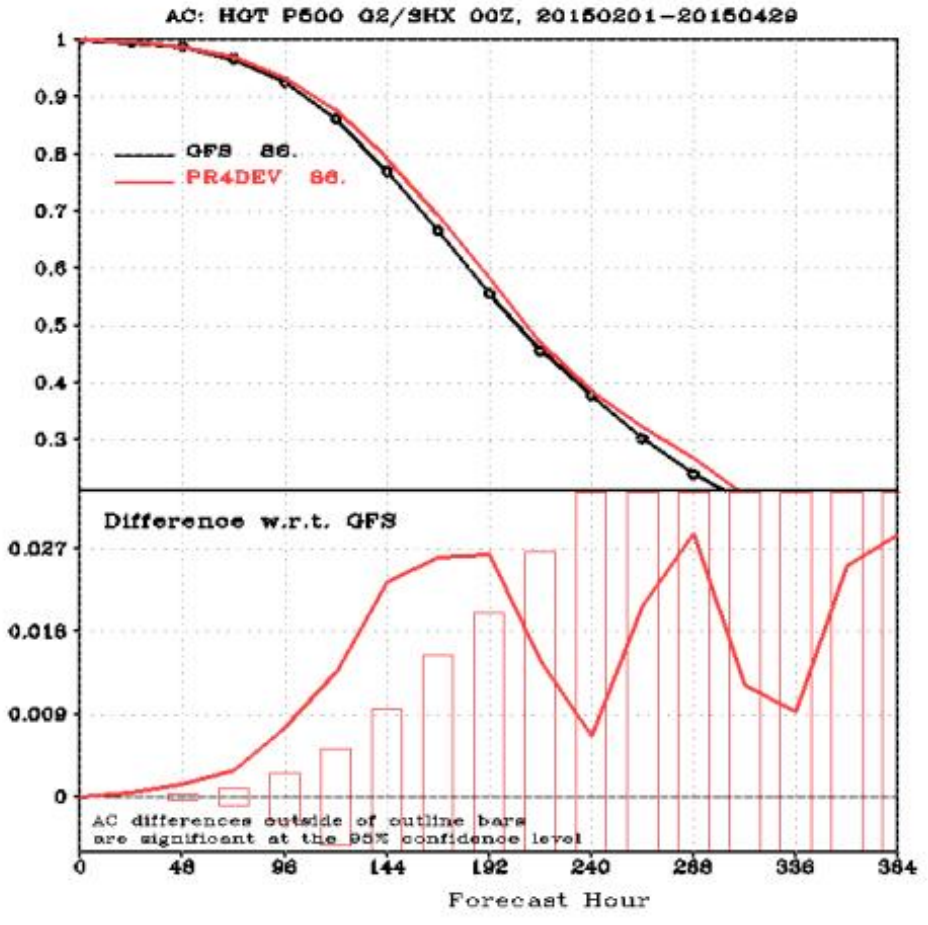



Trial run 3D Hybrid minus 3DVAR for various metrics using the T574/T254 Eulerian GFS configuration, Feb 1 through May 15 2012 [with additive inflation, old tuning]

Trial run 4D Hybrid minus 3D Hybrid [EnVar] for various metrics using the T670/T254 Semi-Lagrangian GFS configuration, December 5, 2013 through January 5, 2014 [stochastic physics, new tuning]

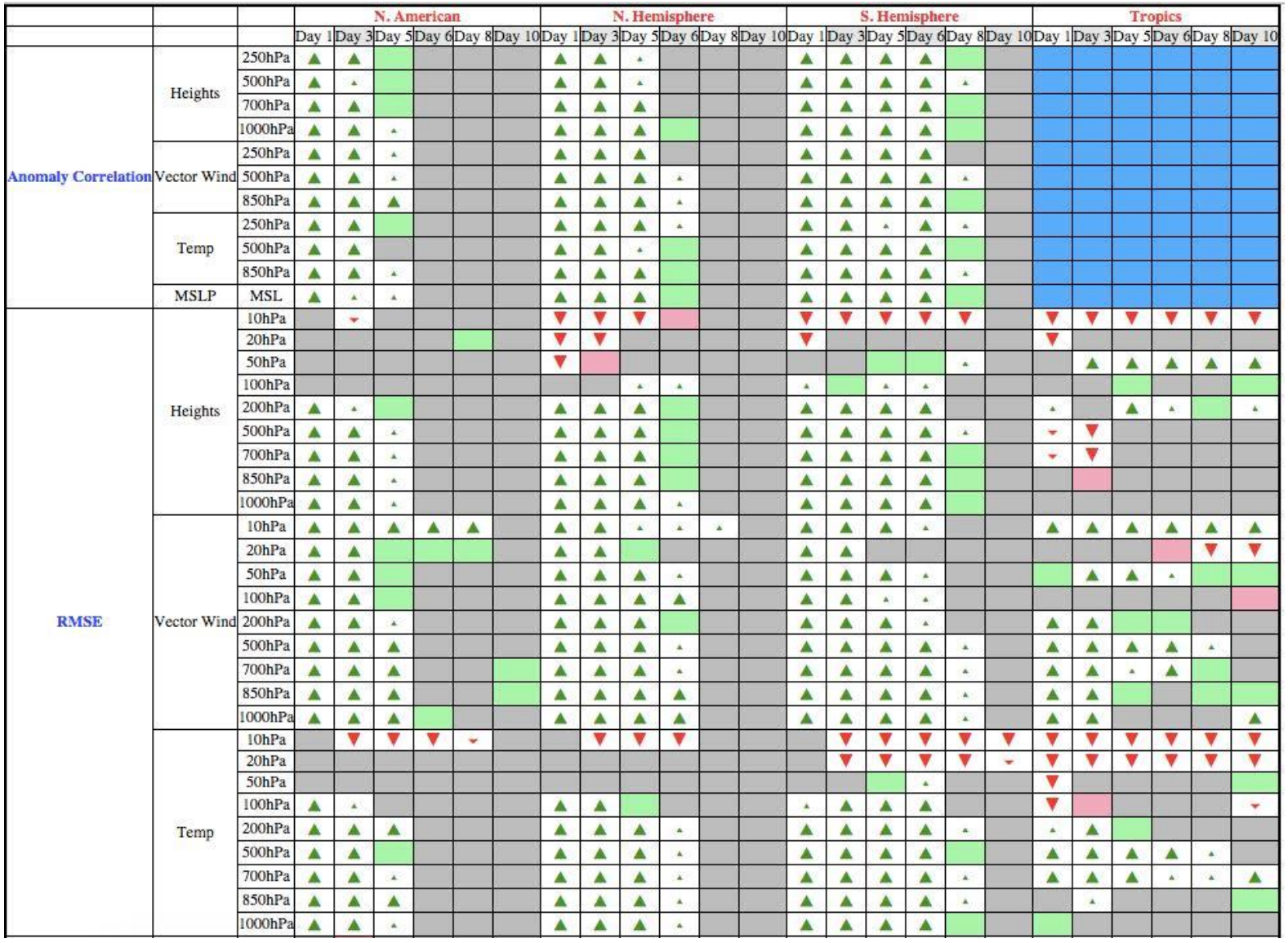
# Full Resolution (T1534/T574) Trials: 500 hPa AC




500 hPa AC for the Operational GFS (Black, 3D Hybrid) and Test 4D configuration (Red) for the period covering 02-01-2015 through 04-29-2015.

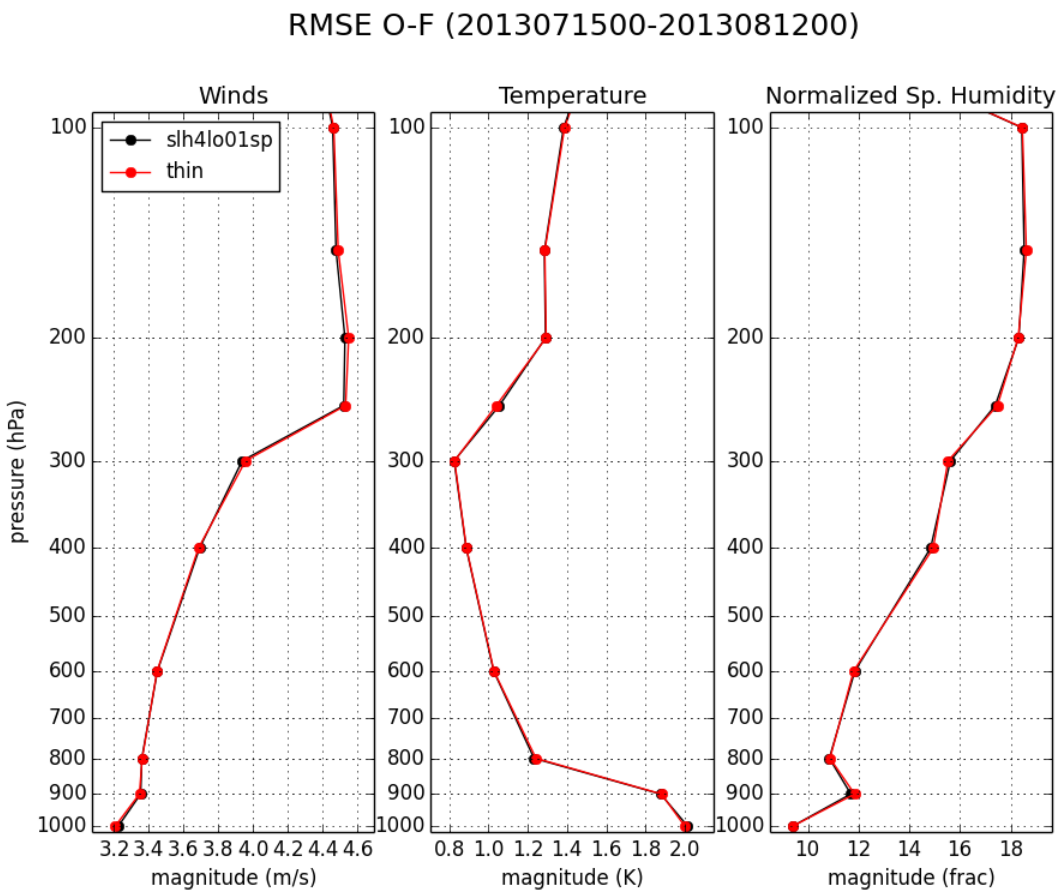
# Full Resolution (T1534/T574) Trials: Summary Scorecard (02-01 through 04-29 2015)



# Hybrid 4D EnVar: Summary

- Natural extension to operational EnVar. No need for development of maintenance of TLM and ADJ models
- Highly scalable
  - And can be improved even further
  - Aligns with technological/computing advances
- Stochastic Physics
  - Suitable replacement for additive inflation.
- Computationally inexpensive relative to 4DVAR (with TL/AD)
  - Estimates of improved efficiency by 10x or more
- Targeted for implementation at NCEP (end 2015/early 2016)
  - Still lots of work to do after implementation
  - 4D EnVar now operational at CMC, being tested at UKMO

# Data Selection in 4D EnVar
# Courtesy: Catherine Thomas




Default data selection in GSI gives preferential treatment to observations in middle of window

In 4D modes (4DVar or 4DEnVar), want more even distribution of observations within window to constraint solution

Relaxation of data selection yields neutral impact (left)

Ongoing work to expand 4D data thinning by actual observation bin

# Potential Corrections for Noise and/or Imbalance

- Noise in the background (first guess/model forecast)
  - Full field digital filters ** (currently used in GFS)
  - Initialization (Nonlinear Normal Mode Initialization)
    - Analysis draws to data, Initialization pushes away from observations

- Noise in the analysis increment
  - Improved multivariate variable definition
  - Penalty terms
  - *Incremental normal mode initialization*

- Discrepancy in passing increment to model
  - Incremental analysis update

# 4D Incremental Analysis Update: Motivation

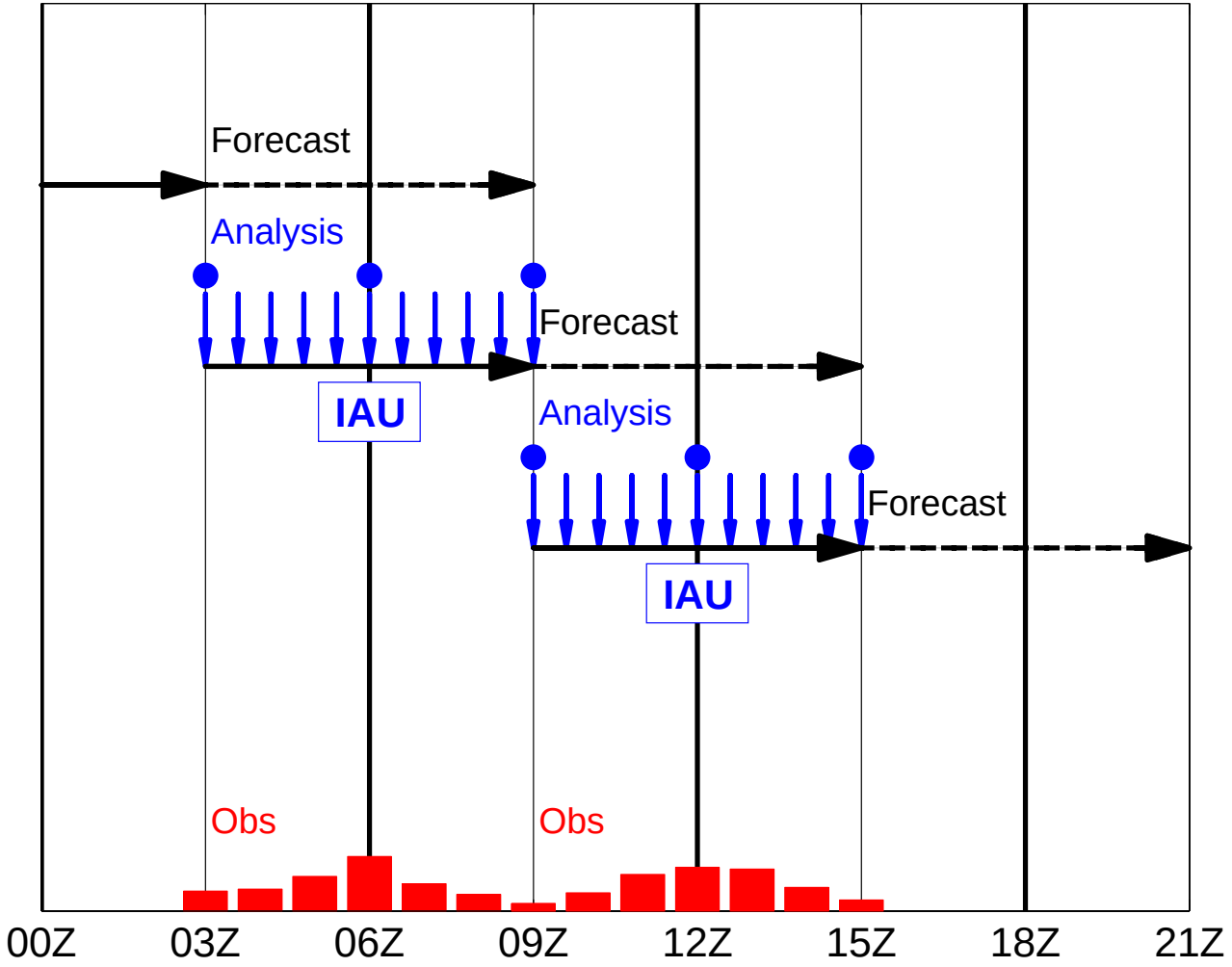
- Imbalances generated by discontinuous nature of analysis, localization & inflation (Greybush, 2011; Kepert, 2009).

- Incremental Analysis Update (Bloom, 1996) helps by using model to distribute a (single) increment over a time window with constant weights (we call this 3DIAU).
  - Propagation of increment neglected, might be significant for fast-moving weather systems.
  - May help spin up unobserved/non-updated state variables

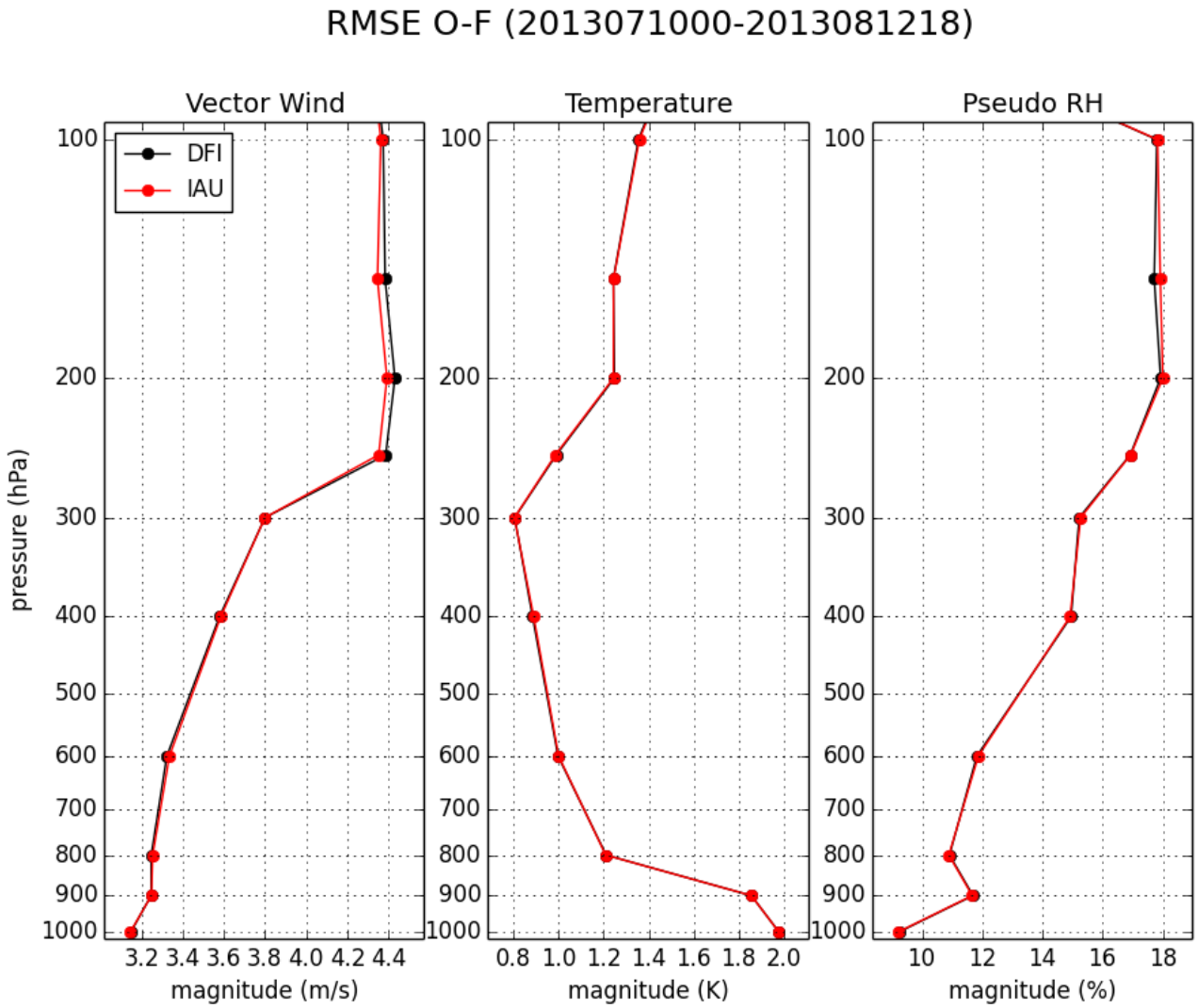
- 4D version of IAU has been proposed by UK Met Office
  - Positive Impact in UKMO and Canadian 4D EnVar

- Approximation of “mollified” time-continuous formulation EnKF proposed by Bergemann & Reich (2010).

# Schematic of 4DIAU
# Courtesy: Jeff Whitaker and Lili Lei



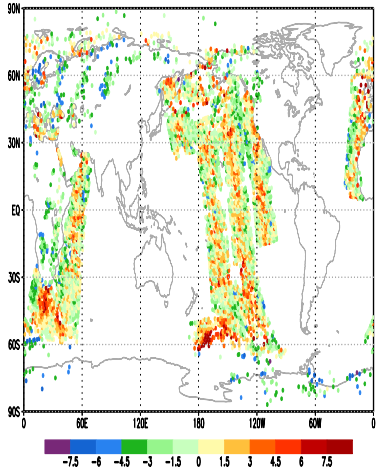
# Preliminary Low Res. Results
# Courtesy: Rahul Mahajan



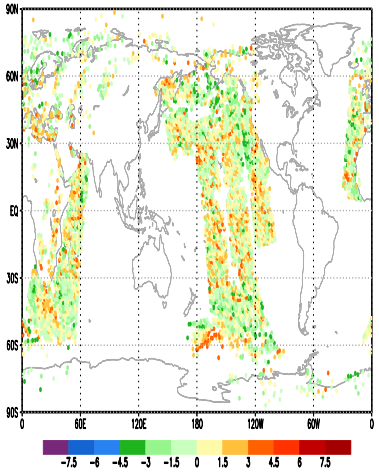
# Other Significant DA Developments besides coupled prototype and EnVar

# Enhanced Radiance Bias Correction Scheme
# Courtesy: Yanqiu Zhu

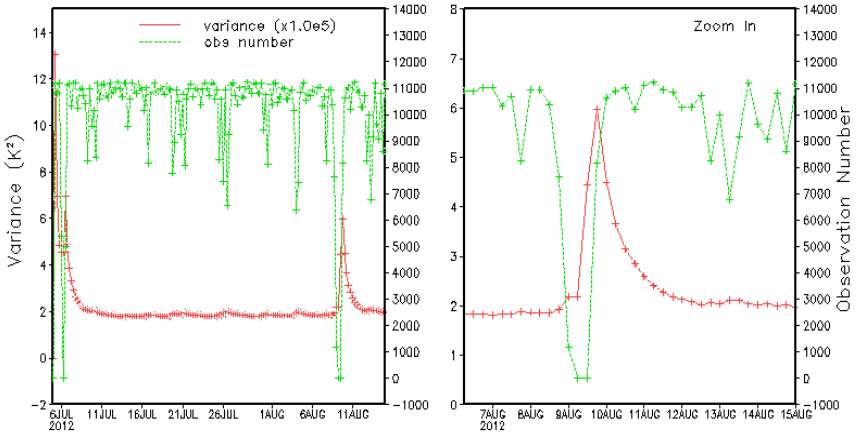
- Simplify operational suite – a single step variational procedure inside the GSI
- Built-in bias initialization step for new radiance data
- Automatically detect any new/missing/recovery of radiance data
- Adaptive background error variance for bias coefficients -- quickly capture any changes in the data and the data assimilation system
- New passive channel bias correction capability -- an efficient way to obtain the bias of any new satellite data that are not used but monitored for preparation for future use, such as the radiance data from the JPSS program



OmF before bias correction



OmF after bias correction

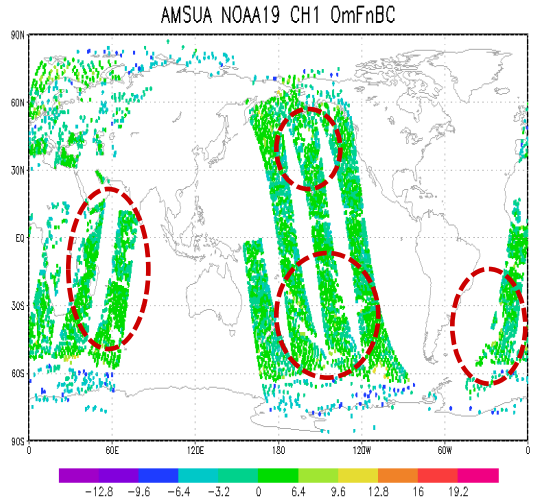



Error variance vs. Observation number

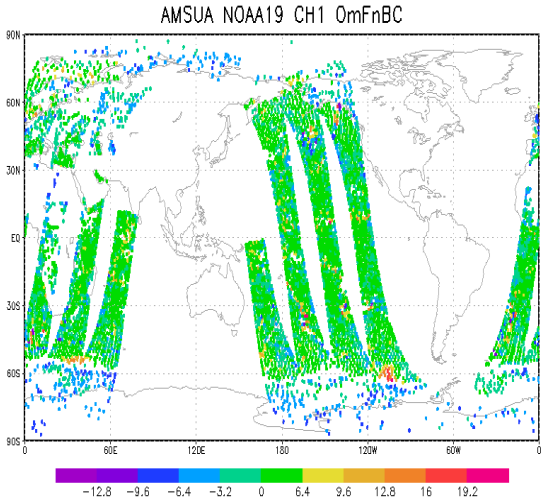
Zhu, Y., J. Derber, A. Collard, D. Dee, R. Treadon, G. Gayno, and J. A. Jung 2014: Enhanced radiance bias correction in the National Centers for Environmental Prediction's Gridpoint Statistical Interpolation Data Assimilation System. Q. J. R. Meteorol. Soc., 140, 1479-1492.

# Cloudy/Precipitating Radiance Assimilation
# Courtesy: Yanqiu Zhu and Emily Liu

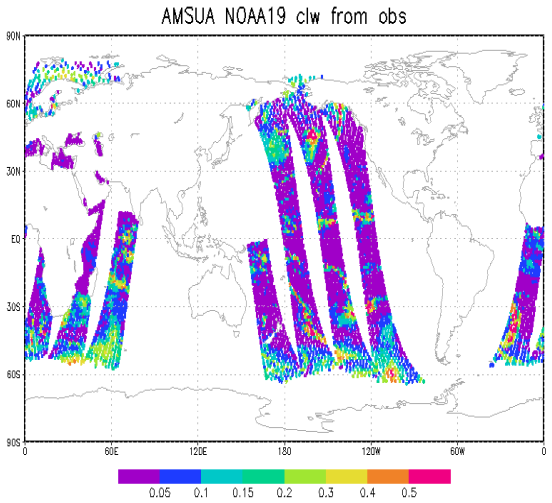
## Clear Sky O-F




## "All Sky" O-F




## CLW

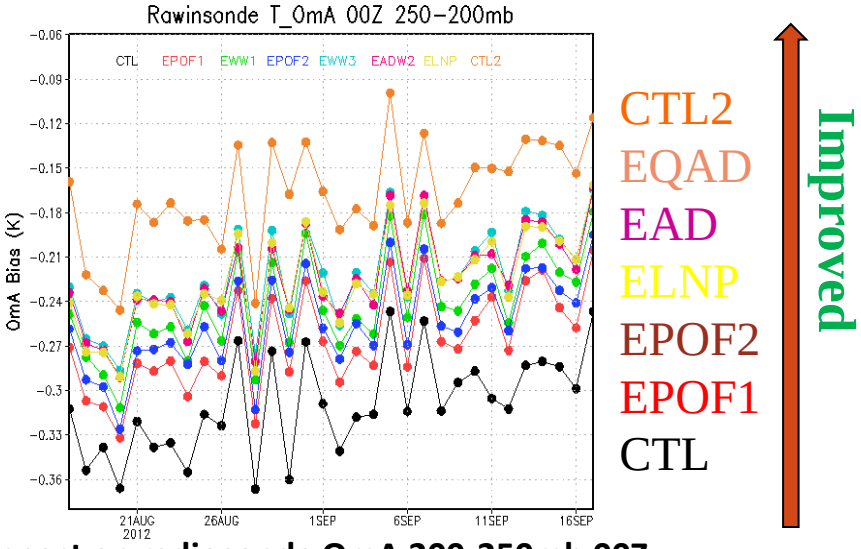



- Thick clouds that are excluded from clear-sky assimilation are now assimilated under all-sky condition
- Heavily precipitating areas are excluded from both conditions

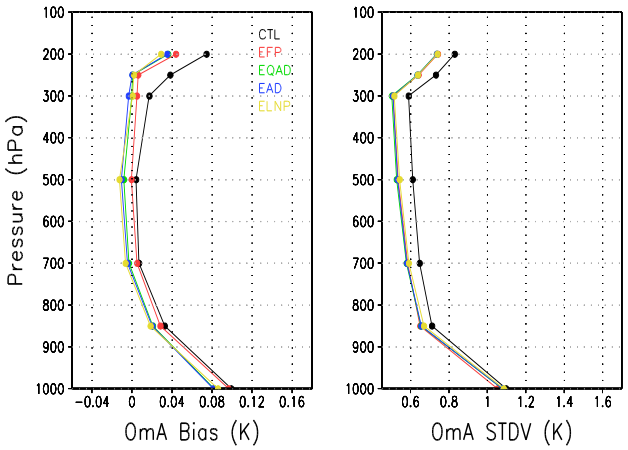
AMSUA NOAA19 CH1 00Z 20131029

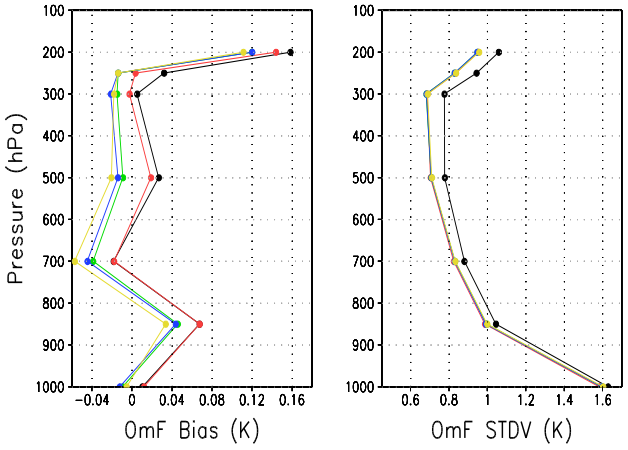
# Variational Bias Correction of Aircraft Temperatures
## Courtesy: Yanqiu Zhu

- Tail-number dependent bias correction.
- Bias models tested: Flight phase dependent, Quadratic aircraft ascent/descent rate, linear separate aircraft ascent/descent rate, logarithm pressure.
- Quadratic aircraft ascent/descent rate bias model + aircraft ascent/descent rate calculated from tensioned-splines algorithm currently is tested in the 4DEnVar




**Impact on radiosonde OmA 200-250mb 00Z**




Aircraft temperature OmF bias and STDV

Zhu, Y., J. Derber, R. J. Purser, B. A. Ballish, and J. Whiting, 2015: Variational correction of aircraft temperature bias in the NCEP's GSI analysis system. Mon. Wea. Rev., in press.

# Summary

- Significant progress has been made on 4D EnVar development and testing for operational NWP at NCEP, set to be implemented later this year for GFS. Much of the current and future work should yield improved and more consistent reanalyses in the future.

- Future work on 4D EnVar DA at NCEP/UMD:
  - Scale-dependent weighting (visitor Deng-shun Chen)
  - Localization: Wave-band/scale-dependent
  - Initialization: Synergy between EnVar and EnKF?
  - Incorporate ensemble update into GSI (EVIL, d-EVIL, mean-pert)
  - Nonlinearity, outer loops, variable choices
  - What to do about static (time-invariant) static error covariance

- Other exciting DA developments relevant to reanalysis: bias correction (radiance and conventional), cloudy/precipitating fields of view, new variables